

# SEQUENCE LIST

SEQ ID NO.: 1

Length: 417 amino acid

Type: amino acid

5 Topology: linear

Molecule type: protein

Sequence

Met Ala Leu Lys Val Leu Leu Glu Gln Glu Lys Thr Phe Phe Thr Leu

1 5 10 15

10 Leu Val Leu Leu Gly Tyr Leu Ser Cys Lys Val Thr Cys Glu Thr Gly

20 25 30

Asp Cys Arg Gln Gln Glu Phe Arg Asp Arg Ser Gly Asn Cys Val Pro

35 40 45

Cys Asn Gln Cys Gly Pro Gly Met Glu Leu Ser Lys Glu Cys Gly Phe

15 50 55 60

Gly Tyr Gly Glu Asp Ala Gln Cys Val Thr Cys Arg Leu His Arg Phe

65 70 75 80

Lys Glu Asp Trp Gly Phe Gln Lys Cys Lys Pro Cys Leu Asp Cys Ala

85 90 95

20 Val Val Asn Arg Phe Gln Lys Ala Asn Cys Ser Ala Thr Ser Asp Ala

100 105 110

Ile Cys Gly Asp Cys Leu Pro Gly Phe Tyr Arg Lys Thr Lys Leu Val

115 120 125

Gly Phe Gln Asp Met Glu Cys Val Pro Cys Gly Asp Pro Pro Pro

25 130 135 140

	Tyr	Glu	Pro	His	Cys	Ala	Ser	Lys	Val	Asn	Leu	Val	Lys	Ile	Ala	Ser
	145					150					155					160
	Thr	Ala	Ser	Ser	Pro	Arg	Asp	Thr	Ala	Leu	Ala	Ala	Val	Ile	Cys	Ser
					165					170					175	
5	Ala	Leu	Ala	Thr	Val	Leu	Leu	Ala	Leu	Leu	Ile	Leu	Cys	Val	Ile	Tyr
		180						185						190		
	Cys	Lys	Arg	Gln	Phe	Met	Glu	Lys	Lys	Pro	Ser	Trp	Ser	Leu	Arg	Ser
		195						200					205			
	Gln	Asp	Ile	Gln	Tyr	Asn	Gly	Ser	Glu	Leu	Ser	Cys	Leu	Asp	Pro	Arg
10		210					215					220				
	Gln	Leu	His	Glu	Tyr	Ala	His	Arg	Ala	Cys	Cys	Gln	Cys	Arg	Arg	Asp
	225					230				235				240		
	Ser	Val	Gln	Thr	Cys	Gly	Pro	Val	Arg	Leu	Leu	Pro	Ser	Met	Cys	Cys
					245				250					255		
15	Glu	Glu	Ala	Cys	Ser	Pro	Asn	Pro	Ala	Thr	Leu	Gly	Cys	Gly	Val	His
		260						265					270			
	Ser	Ala	Ala	Ser	Leu	Gln	Ala	Arg	Asn	Ala	Gly	Pro	Ala	Gly	Glu	Met
		275						280					285			
	Val	Pro	Thr	Phe	Phe	Gly	Ser	Leu	Thr	Gln	Ser	Ile	Cys	Gly	Glu	Phe
20		290					295					300				
	Ser	Asp	Ala	Trp	Pro	Leu	Met	Gln	Asn	Pro	Met	Gly	Gly	Asp	Asn	Ile
	305					310				315				320		
	Ser	Phe	Cys	Asp	Ser	Tyr	Pro	Glu	Leu	Thr	Gly	Glu	Asp	Ile	His	Ser
					325					330				335		
25	Leu	Asn	Pro	Glu	Leu	Glu	Ser	Ser	Thr	Ser	Leu	Asp	Ser	Asn	Ser	Ser

340 345 350  
Gln Asp Leu Val Gly Gly Ala Val Pro Val Gln Ser His Ser Glu Asn

355 360 365  
Phe Thr Ala Ala Thr Asp Leu Ser Arg Tyr Asn Asn Thr Leu Val Glu

5 370 375 380  
Ser Ala Ser Thr Gln Asp Ala Leu Thr Met Arg Ser Gln Leu Asp Gln

385 390 395 400  
Glu Ser Gly Ala Ile Ile His Pro Ala Thr Gln Thr Ser Leu Gln Glu

10 405 410 415  
Ala

SEQ ID NO.: 2

Length: 1269 base pairs

Type: nucleic acid

15 Strandness: single

Topology: linear

Molecule type: cDNA to mRNA

Sequence

20 ATGGCTTTAA AAGTGCTACT AGAACAAGAG AAAACGTTTT TCACTCTTTT AGTATTACTA 60  
GGCTATTTGT CATGTAAAGT GACTTGTGAA ACAGGAGACT GTAGACAGCA AGAATTCAGG 120  
GATCGGTCTG GAAACTGTGT TCCCTGCAAC CAGTGTGGGC CAGGCATGGA GTTGTCTAAG 180  
GAATGTGGCT TCGGCTATGG GGAGGATGCA CAGTGTGTGA CGTGCCGGCT GCACAGGTTC 240  
AAGGAGGACT GGGGCTTCCA GAAATGCAAG CCCTGTCTGG ACTGCGCAGT GGTGAACCGC 300  
25 TTTCAGAAGG CAAATTGTTC AGCCACCAGT GATGCCATCT GCGGGGACTG CTTGCCAGGA 360



GGGAACGTAG AACTCTCCAA CAATAAATAC ATTTGATAAG AAAGATGGCT TTAAAAGTGC 60  
TACTAGAACA AGAGAAAACG TTTTCTACTC TTTTAGTATT ACTAGGCTAT TTGTCATGTA 120  
AAGTGACTTG TGAAACAGGA GACTGTAGAC AGCAAGAATT CAGGGATCGG TCTGGAAACT 180  
GTGTTCCCTG CAACCAGTGT GGGCCAGGCA TGGAGTTGTC TAAGGAATGT GGCTTCGGCT 240  
5 ATGGGGAGGA TGCACAGTGT GTGACGTGCC GGCTGCACAG GTTCAAGGAG GACTGGGGCT 300  
TCCAGAAATG CAAGCCCTGT CTGGACTGCG CAGTGGTGAA CCGCTTTCAG AAGGCAAATT 360  
GTTCAGCCAC CAGTGATGCC ATCTGCGGGG ACTGCTTGCC AGGATTTTAT AGGAAGACGA 420  
AACTTGTCGG CTTTCAAGAC ATGGAGTGTG TGCCTTGTGG AGACCCTCCT CCTCCTTACG 480  
AACCGCACTG TGCCAGCAAG GTCAACCTCG TGAAGATCGC GTCCACGGCC TCCAGCCCAC 540  
10 GGGACACGGC GCTGGCTGCC GTTATCTGCA GCGCTCTGGC CACCGTCCTG CTGGCCCTGC 600  
TCATCCTCTG TGTCATCTAT TGTAAGAGAC AGTTTATGGA GAAGAAACCC AGCTGGTCTC 660  
TGCGGTCACA GGACATTCAG TACAACGGCT CTGAGCTGTC GTGTCTTGAC AGACCTCAGC 720  
TCCACGAATA TGCCACAGA GCCTGCTGCC AGTGCCGCCG TGA CTCACTG CAGACCTGCG 780  
GGCCGGTGCG CTTGCTCCCA TCCATGTGCT GTGAGGAGGC CTGCAGCCCC AACCCGGCGA 840  
15 CTCTTGTTG TGGGGTG CAT TCTGCAGCCA GTCTTCAGGC AAGAAACGCA GGCCAGCCG 900  
GGGAGATGGT GCCGACTTTC TTCGGATCCC TCACGCAGTC CATCTGTGGC GAGTTTTTCTG 960  
ATGCCTGGCC TCTGATGCAG AATCCCATGG GTGGTGACAA CATCTCTTTT TGTGACTCTT 1020  
ATCCTGAACT CACTGGAGAA GACATTCATT CTCTCAATCC AGAACTTGAA AGCTCAACGT 1080  
CTTTGGATTG AAATAGCAGT CAAGATTTGG TTGGTGGGGC TGTTCAGTC CAGTCTCATT 1140  
20 CTGAAAACCTT TACAGCAGCT ACTGATTTAT CTAGATATAA CAACACACTG GTAGAATCAG 1200  
CATCAACTCA GGATGCACTA ACTATGAGAA GCCAGCTAGA TCAGGAGAGT GCGCTATCA 1260  
TCCACCCAGC CACTCAGACG TCCCTCCAGG AAGCTTAAAG AACCTGCTTC TTTCTGCAGT 1320  
AGAAGCGTGT GCTGGAACCC AAAGAGTACT CCTTTGTTAG GCTTATGGAC TGAGCAGTCT 1380  
GGACCTTGCA TGGCTTCTGG GGCAAAAATA AATCTGAACC AACTGACGG CATTTGAAGC 1440  
25 CTTTCAGCCA GTTGCTTCTG AGCCAGACCA GCTGTAAGCT GAAACCTCAA TGAATAACAA 1500

GAAAAGACTC CAGGCCGACT CATGATACTC TGCATCTTTC CTACATGAGA AGCTTCTCTG 1560  
 CCACAAAAGT GACTTCAAAG ACGGATGGGT TGAGCTGGCA GCCTATGAGA TTGTGGACAT 1620  
 ATAACAAGAA ACAGAAATGC CCTCATGCTT ATTTTCATGG TGATTGTGGT TTTACAAGAC 1680  
 TGAAGACCCA GAGTATACTT TTTC 1704

5

SEQ ID NO.: 4

Length: 1704 base pairs

Type: nucleic acid

Strandness: single

10 Topology: linear

Molecule type: cDNA to mRNA

Original source:

Organism: Homo Sapiens

Cell line: HAS303

15 Feature

Name/Key: CDS

Location: 45..1295

Identification method: P

Name/Key: sig peptide

20 Location: 45..119

Identification method: S

Name/Key: mat peptide

Location: 120..1295

Identification method: S

25 Sequecne

GGGAACGTAG AACTCTCCAA CAATAAATAC ATTTGATAAG AAAG ATG GCT TTA AAA 56

Met Ala Leu Lys

-25

5 GTG CTA CTA GAA CAA GAG AAA ACG TTT TTC ACT CTT TTA GTA TTA CTA 104

Val Leu Leu Glu Gln Glu Lys Thr Phe Phe Thr Leu Leu Val Leu Leu

-20

-15

-10

GGC TAT TTG TCA TGT AAA GTG ACT TGT GAA ACA GGA GAC TGT AGA CAG 152

Gly Tyr Leu Ser Cys Lys Val Thr Cys Glu Thr Gly Asp Cys Arg Gln

10 -5 1 5 10

CAA GAA TTC AGG GAT CGG TCT GGA AAC TGT GTT CCC TGC AAC CAG TGT 200

Gln Glu Phe Arg Asp Arg Ser Gly Asn Cys Val Pro Cys Asn Gln Cys

15

20

25

GGG CCA GGC ATG GAG TTG TCT AAG GAA TGT GGC TTC GGC TAT GGG GAG 248

15 Gly Pro Gly Met Glu Leu Ser Lys Glu Cys Gly Phe Gly Tyr Gly Glu

30

35

40

GAT GCA CAG TGT GTG ACG TGC CGG CTG CAC AGG TTC AAG GAG GAC TGG 296

Asp Ala Gln Cys Val Thr Cys Arg Leu His Arg Phe Lys Glu Asp Trp

45

50

55

20 GGC TTC CAG AAA TGC AAG CCC TGT CTG GAC TGC GCA GTG GTG AAC CGC 344

Gly Phe Gln Lys Cys Lys Pro Cys Leu Asp Cys Ala Val Val Asn Arg

60

65

70

75

TTT CAG AAG GCA AAT TGT TCA GCC ACC AGT GAT GCC ATC TGC GGG GAC 392

Phe Gln Lys Ala Asn Cys Ser Ala Thr Ser Asp Ala Ile Cys Gly Asp

25

80

85

90

	TGC TTG CCA GGA TTT TAT AGG AAG ACG AAA CTT GTC GGC TTT CAA GAC	440
	Cys Leu Pro Gly Phe Tyr Arg Lys Thr Lys Leu Val Gly Phe Gln Asp	
	95 100 105	
	ATG GAG TGT GTG CCT TGT GGA GAC CCT CCT CCT CCT TAC GAA CCG CAC	488
5	Met Glu Cys Val Pro Cys Gly Asp Pro Pro Pro Pro Tyr Glu Pro His	
	110 115 120	
	TGT GCC AGC AAG GTC AAC CTC GTG AAG ATC GCG TCC ACG GCC TCC AGC	536
	Cys Ala Ser Lys Val Asn Leu Val Lys Ile Ala Ser Thr Ala Ser Ser	
	125 130 135	
10	CCA CGG GAC ACG GCG CTG GCT GCC GTT ATC TGC AGC GCT CTG GCC ACC	584
	Pro Arg Asp Thr Ala Leu Ala Ala Val Ile Cys Ser Ala Leu Ala Thr	
	140 145 150 155	
	GTC CTG CTG GCC CTG CTC ATC CTC TGT GTC ATC TAT TGT AAG AGA CAG	632
	Val Leu Leu Ala Leu Leu Ile Leu Cys Val Ile Tyr Cys Lys Arg Gln	
15	160 165 170	
	TTT ATG GAG AAG AAA CCC AGC TGG TCT CTG CGG TCA CAG GAC ATT CAG	680
	Phe Met Glu Lys Lys Pro Ser Trp Ser Leu Arg Ser Gln Asp Ile Gln	
	175 180 185	
	TAC AAC GGC TCT GAG CTG TCG TGT CTT GAC AGA CCT CAG CTC CAC GAA	728
20	Tyr Asn Gly Ser Glu Leu Ser Cys Leu Asp Rro Arg Gln Leu His Glu	
	190 195 200	
	TAT GCC CAC AGA GCC TGC TGC CAG TGC CGC CGT GAC TCA GTG CAG ACC	776
	Tyr Ala His Arg Ala Cys Cys Gln Cys Arg Arg Asp Ser Val Gln Thr	
	205 210 215	
25	TGC GGG CCG GTG CGC TTG CTC CCA TCC ATG TGC TGT GAG GAG GCC TGC	824



	Cys Gly Pro Val Arg Leu Leu Pro Ser Met Cys Cys Glu Glu Ala Cys	
	220	225 230 235
	AGC CCC AAC CCG GCG ACT CTT GGT TGT GGG GTG CAT TCT GCA GCC AGT	872
	Ser Pro Asn Pro Ala Thr Leu Gly Cys Gly Val His Ser Ala Ala Ser	
5	240	245 250
	CTT CAG GCA AGA AAC GCA GGC CCA GCC GGG GAG ATG GTG CCG ACT TTC	920
	Leu Gln Ala Arg Asn Ala Gly Pro Ala Gly Glu Met Val Pro Thr Phe	
	255	260 265
	TTC GGA TCC CTC ACG CAG TCC ATC TGT GGC GAG TTT TCA GAT GCC TGG	968
10	Phe Gly Ser Leu Thr Gln Ser Ile Cys Gly Glu Phe Ser Asp Ala Trp	
	270	275 280
	CCT CTG ATG CAG AAT CCC ATG GGT GGT GAC AAC ATC TCT TTT TGT GAC	1016
	Pro Leu Met Gln Asn Pro Met Gly Gly Asp Asn Ile Ser Phe Cys Asp	
	285	290 295
15	TCT TAT CCT GAA CTC ACT GGA GAA GAC ATT CAT TCT CTC AAT CCA GAA	1064
	Ser Tyr Pro Glu Leu Thr Gly Glu Asp Ile His Ser Leu Asn Pro Glu	
	300	305 310 315
	CTT GAA AGC TCA ACG TCT TTG GAT TCA AAT AGC AGT CAA GAT TTG GTT	1112
	Leu Glu Ser Ser Thr Ser Leu Asp Ser Asn Ser Ser Gln Asp Leu Val	
20	320	325 330
	GGT GGG GCT GTT CCA GTC CAG TCT CAT TCT GAA AAC TTT ACA GCA GCT	1160
	Gly Gly Ala Val Pro Val Gln Ser His Ser Glu Asn Phe Thr Ala Ala	
	335	340 345
	ACT GAT TTA TCT AGA TAT AAC AAC ACA CTG GTA GAA TCA GCA TCA ACT	1208
25	Thr Asp Leu Ser Arg Tyr Asn Asn Thr Leu Val Glu Ser Ala Ser Thr	

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25

```

350              355              360
CAG GAT GCA CTA ACT ATG AGA AGC CAG CTA GAT CAG GAG AGT GGC GCT 1256
Gln Asp Ala Leu Thr Met Arg Ser Gln Leu Asp Gln Glu Ser Gly Ala

365              370              375
5  ATC ATC CAC CCA GCC ACT CAG ACG TCC CTC CAG GAA GCT TAAAGAACCT 1305
Ile Ile His Pro Ala Thr Gln Thr Ser Leu Gln Glu Ala

380              385              390
GCTTCTTTCT GCAGTAGAAG CGTGTGCTGG AACCCAAAGA GTACTCCTTT GTTAGGCTTA 1365
TGGACTGAGC AGTCTGGACC TTGCATGGCT TCTGGGGCAA AAATAAATCT GAACCAAACCT 1425
10 GACGGCATTG GAAGCCTTTC AGCCAGTTGC TTCTGAGCCA GACCAGCTGT AAGCTGAAAC 1485
CTCAATGAAT AACAAGAAAA GACTCCAGGC CGACTCATGA TACTCTGCAT CTTTCCTACA 1545
TGAGAAGCTT CTCTGCCACA AAAGTGACTT CAAAGACGGA TGGGTTGAGC TGGCAGCCTA 1605
TGAGATTGTG GACATATAAC AAGAAACAGA AATGCCCTCA TGCTTATTTT CATGGTGATT 1665
GTGGTTTTAC AAGACTGAAG ACCCAGAGTA TACTTTTTTC 1704

```

SEQ ID NO.: 5

Length: 423 amino acids

Type: amino acid

20 Topology: linear

Molecule type: protein

Sequence

Met Ala Leu Lys Val Leu Leu Glu Gln Glu Lys Thr Phe Phe Thr Leu

25      1                      5                      10                      15

Leu Val Leu Leu Gly Tyr Leu Ser Cys Lys Val Thr Cys Glu Thr Gly

20

25

30

Asp Cys Arg Gln Gln Glu Phe Arg Asp Arg Ser Gly Asn Cys Val Pro

35

40

45

5 Cys Asn Gln Cys Gly Pro Gly Met Glu Leu Ser Lys Glu Cys Gly Phe

50

55

60

Gly Tyr Gly Glu Asp Ala Gln Cys Val Thr Cys Arg Leu His Arg Phe

65

70

75

80

Lys Glu Asp Trp Gly Phe Gln Lys Cys Lys Pro Cys Leu Asp Cys Ala

10

85

90

95

Val Val Asn Arg Phe Gln Lys Ala Asn Cys Ser Ala Thr Ser Asp Ala

100

105

110

Ile Cys Gly Asp Cys Leu Pro Gly Phe Tyr Arg Lys Thr Lys Leu Val

115

120

125

15 Gly Phe Gln Asp Met Glu Cys Val Pro Cys Gly Asp Pro Pro Pro Pro

130

135

140

Tyr Glu Pro His Cys Ala Ser Lys Val Asn Leu Val Lys Ile Ala Ser

145

150

155

160

Thr Ala Ser Ser Pro Arg Asp Thr Ala Leu Ala Ala Val Ile Cys Ser

20

165

170

175

Ala Leu Ala Thr Val Leu Leu Ala Leu Leu Ile Leu Cys Val Ile Tyr

180

185

190

Cys Lys Arg Gln Phe Met Glu Lys Lys Pro Ser Trp Ser Leu Arg Ser

195

200

205

25 Gln Asp Ile Gln Tyr Asn Gly Ser Glu Leu Ser Cys Leu Asp Pro Arg

	210	215	220	
	Gln	Leu	His	Glu Tyr Ala His Arg Ala Cys Cys Gln Cys Arg Arg Asp
	225	230	235	240
	Ser	Val	Gln Thr Cys Gly Pro Val Arg Leu Leu Pro Ser Met Cys Cys	
5	245	250	255	
	Glu	Glu	Ala Cys Ser Pro Asn Pro Ala Thr Leu Gly Cys Gly Val His	
	260	265	270	
	Ser	Ala	Ala Ser Leu Gln Ala Arg Asn Ala Gly Pro Ala Gly Glu Met	
	275	280	285	
10	Val	Pro	Thr Phe Phe Gly Ser Leu Thr Gln Ser Ile Cys Gly Glu Phe	
	290	295	300	
	Ser	Asp	Ala Trp Pro Leu Met Gln Asn Pro Met Gly Gly Asp Asn Ile	
	305	310	315	320
	Ser	Phe	Cys Asp Ser Tyr Pro Glu Leu Thr Gly Glu Asp Ile His Ser	
15	325	330	335	
	Leu	Asn	Pro Glu Leu Glu Ser Ser Thr Ser Leu Asp Ser Asn Ser Ser	
	340	345	350	
	Gln	Asp	Leu Val Gly Gly Ala Val Pro Val Gln Ser His Ser Glu Asn	
	355	360	365	
20	Phe	Thr	Ala Ala Thr Asp Leu Ser Arg Tyr Asn Asn Thr Leu Val Glu	
	370	375	380	
	Ser	Ala	Ser Thr Gln Asp Ala Leu Thr Met Arg Ser Gln Leu Asp Gln	
	385	390	395	400
	Glu	Ser	Gly Ala Ile Ile His Pro Ala Thr Gln Thr Ser Leu Gln Val	
25	405	410	415	

Arg Gln Arg Leu Gly Ser Leu

420

SEQ ID NO.: 6

5 Length: 1269 base pairs

Type: nucleic acid

Strandness: single

Topology: linear

Molecule type: cDNA to mRNA

10 Sequecne

ATGGCTTTAA AAGTGCTACT AGAACAAGAG AAAACGTTTT TCACTCTTTT AGTATTACTA 60  
GGCTATTTGT CATGTAAAGT GACTTGTGAA ACAGGAGACT GTAGACAGCA AGAATTCAGG 120  
GATCGGTCTG GAAACTGTGT TCCCTGCAAC CAGTGTGGGC CAGGCATGGA GTTGTCTAAG 180  
15 GAATGTGGCT TCGGCTATGG GGAGGATGCA CAGTGTGTGA CGTGCCGGCT GCACAGGTTC 240  
AAGGAGGACT GGGGCTTCCA GAAATGCAAG CCCTGTCTGG ACTGCGCAGT GGTGAACCGC 300  
TTTCAGAAGG CAAATTGTTC AGCCACCAGT GATGCCATCT GCGGGGACTG CTTGCCAGGA 360  
TTTTATAGGA AGACGAAACT TGTCGGCTTT CAAGACATGG AGTGTGTGCC TTGTGGAGAC 420  
CCTCCTCCTC CTTACGAACC GCACTGTGCC AGCAAGGTCA ACCTCGTGAA GATCGCGTCC 480  
20 ACGGCCTCCA GCCCACGGGA CACGGCGCTG GCTGCCGTTA TCTGCAGCGC TCTGGCCACC 540  
GTCCTGCTGG CCCTGCTCAT CCTCTGTGTC ATCTATTGTA AGAGACAGTT TATGGAGAAG 600  
AAACCCAGCT GGTCTCTGCG GTCACAGGAC ATTCAGTACA ACGGCTCTGA GCTGTCGTGT 660  
CTTGACAGAC CTCAGCTCCA CGAATATGCC CACAGAGCCT GCTGCCAGTG CCGCCGTGAC 720  
TCAGTGCAGA CCTGCGGGCC GGTGCGCTTG CTCCCATCCA TGTGCTGTGA GGAGGCCTGC 780  
25 AGCCCCAACC CGGCGACTCT TGGTTGTGGG GTGCATTCTG CAGCCAGTCT TCAGGCAAGA 840

AACGCAGGCC CAGCCGGGGA GATGGTGCCG ACTTTCTTCG GATCCCTCAC GCAGTCCATC 900  
 TGTGGCGAGT TTTCAGATGC CTGGCCTCTG ATGCAGAATC CCATGGGTGG TGACAACATC 960  
 TCTTTTTGTG ACTCTTATCC TGAAGTCACT GGAGAAGACA TTCATTCTCT CAATCCAGAA 1020  
 CTTGAAAGCT CAACGTCTTT GGATTCAAAT AGCAGTCAAG ATTTGGTTGG TGGGGCTGTT 1080  
 5 CCAGTCCAGT CTCATTCTGA AACTTTTACA GCAGCTACTG ATTTATCTAG ATATAACAAC 1140  
 AACTGGTAG AATCAGCATC AACTCAGGAT GCACTAACTA TGAGAAGCCA GCTAGATCAG 1200  
 GAGAGTGGCG CTATCATCCA CCCAGCCACT CAGACGTCCC TCCAGGTAAG GCAGCGACTG 1260  
 GGTTCCTG 1269

10 SEQ ID NO.: 7

Length: 1496 base pairs

Type: nucleic acid

Strandness: single

Topology: linear

15 Molecule type: cDNA to mRNA

Sequence

GGGAACGTAG AACTCTCCAA CAATAAATAC ATTTGATAAG AAAGATGGCT TTAAAAGTGC 60  
 TACTAGAACA AGAGAAAACG TTTTCTACTC TTTTAGTATT ACTAGGCTAT TTGTCATGTA 120  
 20 AAGTGAAGTG TGAAACAGGA GACTGTAGAC AGCAAGAATT CAGGGATCGG TCTGGAAACT 180  
 GTGTTCCCTG CAACCAAGTG GGGCCAGGCA TGGAGTTGTC TAAGGAATGT GGCTTCGGCT 240  
 ATGGGGAGGA TGCACAGTGT GTGACGTGCC GGCTGCACAG GTTCAAGGAG GACTGGGGCT 300  
 TCCAGAAATG CAAGCCCTGT CTGGACTGCG CAGTGGTGAA CCGCTTTCAG AAGGCAAATT 360  
 GTTCAGCCAC CAGTGATGCC ATCTGCGGGG ACTGCTTGCC AGGATTTTAT AGGAAGACGA 420  
 25 AACTTGTCGG CTTTCAAGAC ATGGAGTGTG TGCCTTGTGG AGACCCTCCT CCTCCTTACG 480

AACCGCACTG TGCCAGCAAG GTCAACCTCG TGAAGATCGC GTCCACGGCC TCCAGCCCAC 540  
GGGACACGGC GCTGGCTGCC GTTATCTGCA GCGCTCTGGC CACCGTCCTG CTGGCCCTGC 600  
TCATCCTCTG TGTCATCTAT TGTAAGAGAC AGTTTATGGA GAAGAAACCC AGCTGGTCTC 660  
TGCGGTCACA GGACATTGAG TACAACGGCT CTGAGCTGTC GTGTCTTGAC AGACCTCAGC 720  
5 TCCACGAATA TGCCACAGA GCCTGCTGCC AGTGCCGCCG TGAATCAGTG CAGACCTGCG 780  
GGCCGGTGCG CTTGCTCCCA TCCATGTGCT GTGAGGAGGC CTGCAGCCCC AACCCGGCGA 840  
CTCTTGTTG TGGGGTGAT TCTGCAGCCA GTCTTCAGGC AAGAAACGCA GGCCAGCCG 900  
GGGAGATGGT GCCGACTTTC TTCGGATCCC TCACGCAGTC CATCTGTGGC GAGTTTTTCAG 960  
ATGCCTGGCC TCTGATGCAG AATCCCATGG GTGGTGACAA CATCTCTTTT TGTGACTCTT 1020  
10 ATCTGAACT CACTGGAGAA GACATTCATT CTCTCAATCC AGAACTTGAA AGCTCAACGT 1080  
CTTTGGATTC AAATAGCAGT CAAGATTGCG TTGGTGGGGC TGTTCAGTC CAGTCTCATT 1140  
CTGAAAACCT TACAGCAGCT ACTGATTAT CTAGATATAA CAACACACTG GTAGAATCAG 1200  
CATCAACTCA GGATGCACTA ACTATGAGAA GCCAGCTAGA TCAGGAGAGT GGCCTATCA 1260  
TCCACCCAGC CACTCAGACG TCCCTCCAGG TAAGGCAGCG ACTGGGTTCC CTGTGAACAC 1320  
15 AGCACTGACT TACAGTAGAT CAGAACTCTG TTCCAGCAT AAGATTTGGG GGAACCTGAT 1380  
GAGTTTTTTT TTTGCATCTT TAATAATTTC TTGTATGTTG TAGAGTATGT TTAAAATAA 1440  
ATTTCAAGTA TTTTTTTTAA AACTAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAA 1496

SEQ ID NO.: 8

20 Length: 1496 base pairs

Type: nucleic acid

Strandness: single

Topology: linear

Molecule type: cDNA to mRNA

25

Original source

Organism: Homo Sapiens

Cell line: HAS303

Feature

5 Name/Key: CDS  
Location: 45..1313  
Identification method: P

10 Name/Key: sig peptide  
Location: 45..119  
Identification method: S

Name/Key: mat peptide  
Location: 120..1313  
Identification method: S

15 Sequence

GGGAACGTAG AACTCTCCAA CAATAAATAC ATTTGATAAG AAAG ATG GCT TTA AAA 56

Met Ala Leu Lys

-25

20 GTG CTA CTA GAA CAA GAG AAA ACG TTT TTC ACT CTT TTA GTA TTA CTA 104

Val Leu Leu Glu Gln Glu Lys Thr Phe Phe Thr Leu Leu Val Leu Leu

-20

-15

-10

GGC TAT TTG TCA TGT AAA GTG ACT TGT GAA ACA GGA GAC TGT AGA CAG 152

Gly Tyr Leu Ser Cys Lys Val Thr Cys Glu Thr Gly Asp Cys Arg Gln

25 -5 1 5 10



	CAA GAA TTC AGG GAT CGG TCT GGA AAC TGT GTT CCC TGC AAC CAG TGT	200
	Gln Glu Phe Arg Asp Arg Ser Gly Asn Cys Val Pro Cys Asn Gln Cys	
	15 20 25	
	GGG CCA GGC ATG GAG TTG TCT AAG GAA TGT GGC TTC GGC TAT GGG GAG	248
5	Gly Pro Gly Met Glu Leu Ser Lys Glu Cys Gly Phe Gly Tyr Gly Glu	
	30 35 40	
	GAT GCA CAG TGT GTG ACG TGC CGG CTG CAC AGG TTC AAG GAG GAC TGG	296
	Asp Ala Gln Cys Val Thr Cys Arg Leu His Arg Phe Lys Glu Asp Trp	
	45 50 55	
10	GGC TTC CAG AAA TGC AAG CCC TGT CTG GAC TGC GCA GTG GTG AAC CGC	344
	Gly Phe Gln Lys Cys Lys Pro Cys Leu Asp Cys Ala Val Val Asn Arg	
	60 65 70 75	
	TTT CAG AAG GCA AAT TGT TCA GCC ACC AGT GAT GCC ATC TGC GGG GAC	392
	Phe Gln Lys Ala Asn Cys Ser Ala Thr Ser Asp Ala Ile Cys Gly Asp	
15	80 85 90	
	TGC TTG CCA GGA TTT TAT AGG AAG ACG AAA CTT GTC GGC TTT CAA GAC	440
	Cys Leu Pro Gly Phe Tyr Arg Lys Thr Lys Leu Val Gly Phe Gln Asp	
	95 100 105	
	ATG GAG TGT GTG CCT TGT GGA GAC CCT CCT CCT CCT TAC GAA CCG CAC	488
20	Met Glu Cys Val Pro Cys Gly Asp Pro Pro Pro Pro Tyr Glu Pro His	
	110 115 120	
	TGT GCC AGC AAG GTC AAC CTC GTG AAG ATC GCG TCC ACG GCC TCC AGC	536
	Cys Ala Ser Lys Val Asn Leu Val Lys Ile Ala Ser Thr Ala Ser Ser	
	125 130 135	
25	CCA CGG GAC ACG GCG CTG GCT GCC GTT ATC TGC AGC GCT CTG GCC ACC	584

	Pro Arg Asp Thr Ala Leu Ala Ala Val Ile Cys Ser Ala Leu Ala Thr	
	140                      145                      150                      155	
	GTC CTG CTG GCC CTG CTC ATC CTC TGT GTC ATC TAT TGT AAG AGA CAG	632
	Val Leu Leu Ala Leu Leu Ile Leu Cys Val Ile Tyr Cys Lys Arg Gln	
5	160                      165                      170	
	TTT ATG GAG AAG AAA CCC AGC TGG TCT CTG CGG TCA CAG GAC ATT CAG	680
	Phe Met Glu Lys Lys Pro Ser Trp Ser Leu Arg Ser Gln Asp Ile Gln	
	175                      180                      185	
	TAC AAC GGC TCT GAG CTG TCG TGT CTT GAC AGA CCT CAG CTC CAC GAA	728
10	Tyr Asn Gly Ser Glu Leu Ser Cys Leu Asp Rro Arg Gln Leu His Glu	
	190                      195                      200	
	TAT GCC CAC AGA GCC TGC TGC CAG TGC CGC CGT GAC TCA GTG CAG ACC	776
	Tyr Ala His Arg Ala Cys Cys Gln Cys Arg Arg Asp Ser Val Gln Thr	
	205                      210                      215	
15	TGC GGG CCG GTG CGC TTG CTC CCA TCC ATG TGC TGT GAG GAG GCC TGC	824
	Cys Gly Pro Val Arg Leu Leu Pro Ser Met Cys Cys Glu Glu Ala Cys	
	220                      225                      230                      235	
	AGC CCC AAC CCG GCG ACT CTT GGT TGT GGG GTG CAT TCT GCA GCC AGT	872
	Ser Pro Asn Pro Ala Thr Leu Gly Cys Gly Val His Ser Ala Ala Ser	
20	240                      245                      250	
	CTT CAG GCA AGA AAC GCA GGC CCA GCC GGG GAG ATG GTG CCG ACT TTC	920
	Leu Gln Ala Arg Asn Ala Gly Pro Ala Gly Glu Met Val Pro Thr Phe	
	255                      260                      265	
	TTC GGA TCC CTC ACG CAG TCC ATC TGT GGC GAG TTT TCA GAT GCC TGG	968
25	Phe Gly Ser Leu Thr Gln Ser Ile Cys Gly Glu Phe Ser Asp Ala Trp	

	270	275	280	
	CCT CTG ATG CAG AAT CCC ATG GGT GGT GAC AAC ATC TCT TTT TGT GAC			1016
	Pro Leu Met Gln Asn Pro Met Gly Gly Asp Asn Ile Ser Phe Cys Asp			
	285	290	295	
5	TCT TAT CCT GAA CTC ACT GGA GAA GAC ATT CAT TCT CTC AAT CCA GAA			1064
	Ser Tyr Pro Glu Leu Thr Gly Glu Asp Ile His Ser Leu Asn Pro Glu			
	300	305	310	315
	CTT GAA AGC TCA ACG TCT TTG GAT TCA AAT AGC AGT CAA GAT TTG GTT			1112
	Leu Glu Ser Ser Thr Ser Leu Asp Ser Asn Ser Ser Gln Asp Leu Val			
10	320	325	330	
	GGT GGG GCT GTT CCA GTC CAG TCT CAT TCT GAA AAC TTT ACA GCA GCT			1160
	Gly Gly Ala Val Pro Val Gln Ser His Ser Glu Asn Phe Thr Ala Ala			
	335	340	345	
	ACT GAT TTA TCT AGA TAT AAC AAC ACA CTG GTA GAA TCA GCA TCA ACT			1208
15	Thr Asp Leu Ser Arg Tyr Asn Asn Thr Leu Val Glu Ser Ala Ser Thr			
	350	355	360	
	CAG GAT GCA CTA ACT ATG AGA AGC CAG CTA GAT CAG GAG AGT GGC GCT			1256
	Gln Asp Ala Leu Thr Met Arg Ser Gln Leu Asp Gln Glu Ser Gly Ala			
	365	370	375	
20	ATC ATC CAC CCA GCC ACT CAG ACG TCC CTC CAG GTA AGG CAG CGA CTG			1304
	Ile Ile His Pro Ala Thr Gln Thr Ser Leu Gln Val Arg Gln Arg Leu			
	380	385	390	395
	GGT TCC CTG TGAACACAG CACTGACTTA CAGTAGATCA GAACTCTGTT CCCAGCATAA			1362
	Gly Ser Leu			
25	GATTTGGGGG AACCTGATGA GTTTTTTTTT TGCATCTTTA ATAATTTCTT GTATGTTGTA			1422

GAGTATGTTT TAAAATAAAT TTCAAGTATT TTTTAAATA ACTAAAAAA AAAAAAAAAA 1482  
AAAAAAAAA AAAA 1496

SEQ ID NO.: 9

5 Length: 35 base pairs

Type: nucleic acid

Strandness: single

Topology: linear

Sequence

10

CGATTGAATT CTAGACCTGC CTCGAGNNNN NNNNN

SEQ ID NO.: 10

Length: 28 base pairs

15 Type: nucleic acid

Strandness: single

Topology: linear

Sequence

AGAAAGATGG CTTTAAAAGT GCTACTAG

See a1  
Concluded